AI के नैतिक निहितार्थों का अन्वेषण

यह एडिटोरियल 24/08/2023 को 'द हिंदू' में प्रकाशित "Can AI be Ethical and Moral?" लेख पर आधारित है। इसमें मशीनों से संबद्ध प्रोग्रामिंग नैतिकता की जटिलता और AI के उपयोग में विश्व द्वारा सतर्कता रखे जाने की आवश्यकता के बारे में चर्चा की गई है।

प्रिलिम्स के लिये:

कृत्रिम बुद्धिमत्ता (AI), वीक AI, स्ट्रॉन्ग AI, मशीन लर्निंग (ML), डीप लर्निंग (DL)।

मेन्स के लिये:

AI की नैतिक चुनौतियाँ, AI की नैतिक चिंताएँ, आर्टिफिशियल मोरल एजेंट (AMA)

निर्णय प्रक्रिया में, **विशेष रूप से शासन (governance) के मामले में, मानवों की सहायता के लिये मशीनों और** कृत्रिम बुद्धिमत्ता **(Artificial Intelligence- AI) का उपयोग दिनानुदिन बढ़ता ही जा रहा है।** नतीजतन, कई देश अब AI विनियमन लागू कर रहे हैं। सरकारी एजेंसियाँ और नीति निर्माता जटिल पैटर्न का विश्लेषण करने, भविष्य के परिदृश्यों का पूर्वानुमान लगाने और अधिक सूचना-संपन्न अनुशंसाएँ प्रदान करने के लिये AI-संचालित टूल्स का लाभ उठा रहे हैं।

हालाँकि, **निर्णय लेने में AI के उपयोग के साथ कई चुनौतियाँ भी जुड़ी हुई हैं**। AI में इसके द्वारा लर्न किये गए डेटा या इसके क्रिएटर्स के दृष्टिकोण से प्रभावित अंतर्निहित पूर्वाग्रह भी हो सकते हैं। इसके **परिणामस्वरूप अनुचित परिणाम उत्पन्न हो सकते हैं, जो शासन में AI का प्रभावी ढंग से उपयोग करने में एक महत्त्वपूर्ण बाधा उत्पन्न कर सकते हैं।**

कृत्रिम बुद्धिमत्ता (AI) क्या है?

- **परिचय:**
  - **AI किसी कंप्यूटर या कंप्यूटर द्वारा नियंत्रित रोबोट की उन कार्यों को कर सकने की क्षमता है जो आम तौर पर मानवों द्वारा किये जाते हैं**, क्योंकि उनके लिये मानव बुद्धि एवं विवेक की आवश्यकता होती है।
    - हालाँकि **ऐसा कोई AI नहीं है जो एक सामान्य मानव द्वारा किये जाने वाले विभिन्न प्रकार के कार्यों को कर सके**, लेकिन कुछ AI कुछ विशिष्ट कार्यों में मानवों की बराबरी कर सकते हैं।
- **विशेषताएँ एवं घटक**:
  - AI की आदर्श विशेषता है इसकी तर्कसंगतता और ऐसी कार्रवाई कर सकने की क्षमता जिससे **किसी विशिष्ट लक्ष्य को प्राप्त करने का सर्वोत्तम अवसर** प्राप्त होता है। मशीन लर्निंग (Machine Learning- ML) AI का एक सबसेट या उपसमूह है।
    - डीप लर्निंग तकनीक (Deep Learning- DL) टेक्स्ट, इमेज या वीडियो जैसे **बड़ी मात्रा में असंरचित डेटा के अवशोषण के माध्यम से इस स्वचालित लर्निंग को सक्षम बनाती है।**
- **विभिन्न श्रेणियाँ:**
  - वीक AI/ नैरो AI
  - स्ट्रॉन्ग AI

AI कुछ दार्शनिक विचारों से कैसे संबंधित है?

- **कांटवादी नैतिक दर्शन:**
  - इमैनुएल कांट (Immanuel Kant) का नैतिक दर्शन तीन प्रमुख सिद्धांतों पर बल देता है:
    - **स्वायत्तता (Autonomy)**—स्वयं का निर्णय ले सकने की क्षमता,

        - **तर्कसंगतता (Rationality)**—विकल्प के चयन के लिये तर्क एवं कारण का उपयोग करना, और
        - **नैतिक कर्तव्य (Moral Duty)**—नैतिक दायित्वों का पालन करना।
    - **शासन में AI का अनुप्रयोग**: निर्णय लेने की प्रक्रियाओं को AI प्रणालियों को सौंपने का कार्य सूक्ष्म नैतिक तर्क की क्षमता को नष्ट करने का जोखिम रखता है। मानवों के बजाय मशीनों को निर्णयन का कार्य सौंपना कांटवादी नैतिकता के महत्त्वपूर्ण विचारों को कमज़ोर कर सकता है।
- **आबद्ध नैतिकता:**
    - वर्ष 2022 में दो शोधकर्त्ताओं ने **डेल्फी (Delphi)**—जो मानव नैतिक निर्णयों के मॉडलिंग के लिये एक प्रोटोटाइप है, का उपयोग कर **आबद्ध नैतिकता (Bounded Ethicality)** पर शोध किया। उन्होंने पाया कि डेल्फी जैसी मशीनें अनैतिक रूप से कार्य कर सकती हैं यदि परिदृश्य इस तरह से तैयार किया गया हो जो नैतिकता को स्वयं कार्यकरण से पृथक रखता हो।
        - इससे पता चलता है कि आबद्ध नैतिकता का मशीनी संस्करण वैसे ही कार्य करता है जैसे मानव कई बार बिना ग्लानि अनुभव किये और प्रायः **औचित्य (justification)** का उपयोग करते हुए अपनी नैतिकता के विरुद्ध कार्य करते हैं।

नोट: आबद्ध नैतिकता लोगों की नैतिक विकल्प चुनने की क्षमता है जो आंतरिक एवं बाह्य दबावों के कारण प्रायः सीमित या प्रतिबंधित होती है।

- **असिमोव के 'रोबोटिक्स के तीन नियम' के समानांतर:**
    - **असिमोव (Issac Asimov) के 'रोबोटिक्स के तीन नियम'** (Three Laws of Robotics) रोबोटों को नैतिक रूप से व्यवहार करने हेतु मार्गदर्शन प्रदान करने के लिये प्रस्तुत किये गए थे। हालाँकि, असिमोव के काल्पनिक परिदृश्यों में इन नियमों के परिणामस्वरूप **प्रायः अप्रत्याशित और विरोधाभासी परिणाम सामने आये, जो नैतिक रूप से कार्य करने के लिये डिज़ाइन की गई मशीनों में भी नैतिक निर्णय लेने की जटिलता को प्रदर्शित कर रहे थे।**
- **कांट और असिमोव का अंतर्संबंध:**
    - तर्कसंगत नैतिक एजेंसी (Rational Moral Agency) पर कांट का ज़ोर और रोबोट के लिये नैतिक दिशानिर्देशों की असिमोव की काल्पनिक खोज आपस में संबद्ध हैं। यह संयोजन उन नैतिक कठिनाइयों और जटिलताओं को प्रकट करने का कार्य करता है जो तब उत्पन्न होती हैं जब मानवीय ज़िम्मेदारियाँ और कार्य कृत्रिम संस्थाओं को सौंप दिये जाते हैं।

असिमोव के नियम:

- प्रथम नियम: एक रोबोट किसी मानव (human being) को आघात नहीं पहुँचाए अथवा अक्रियता के माध्यम से किसी मानव को आघात पहुँचने की अनुमति नहीं दे;
- द्वितीय नियम: एक रोबोट को मानवों द्वारा दिए गए आदेशों का पालन करना चाहिये, सिवाय इसके कि ऐसे आदेश प्रथम नियम के साथ टकराव की स्थिति में न हों;
- तृतीय नियम: एक रोबोट को अपने अस्तित्व की रक्षा तब तक करनी चाहिये जब तक कि ऐसी रक्षा प्रथम या द्वितीय नियम के साथ टकराव की स्थिति में न हो।
- असिमोव ने बाद में एक और नियम जोड़ा, **जिसे चतुर्थ या शून्यवाँ (zeroth) नियम** कहा जाता है, जो अन्य तीनों नियमों पर अधिभावी है। इसमें कहा गया है कि **"एक रोबोट मानवता (Humanity) को आघात नहीं पहुँचाए अथवा अक्रियता से मानवता को आघात पहुँचने की अनुमति नहीं दे।"**

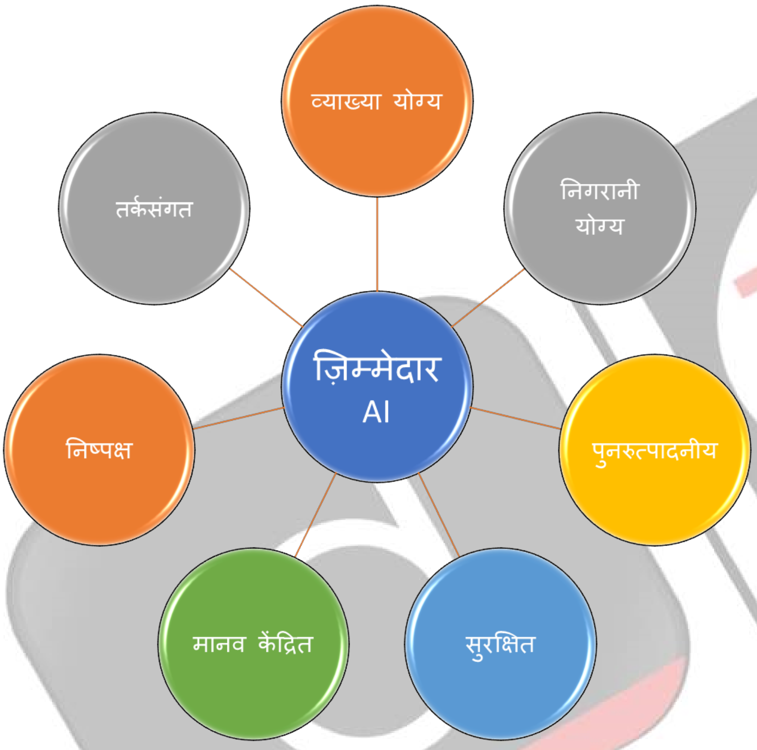
AI से संबद्ध नैतिक चुनौतियाँ

- **रोज़गार विस्थापन और सामाजिक-आर्थिक प्रभाव**: AI द्वारा संचालित स्वचालन (Automation) से कुछ उद्योगों में रोज़गार विस्थापन की स्थिति बन सकती है। बेरोज़गारी और आय असमानता सहित **परिणामी सामाजिक-आर्थिक प्रभाव, इन परिणामों को संबोधित कर सकने में सरकारों और संगठनों की ज़िम्मेदारियों के बारे में नैतिक प्रश्न खड़ा करता है।**
- **नैतिक तर्क के लिये खतरा**: जब परंपरागत रूप से मनुष्यों द्वारा लिये जाने वाले निर्णय एल्गोरिदम और AI को सौंप दिए जाते हैं तो **इससे जोखिम उत्पन्न होता है कि नैतिक तर्क की क्षमता कमज़ोर पड़ेगी**। इसका तात्पर्य यह है कि केवल AI पर निर्भर रहने से **विचारशील नैतिक मनन में संलग्न होने की मानवीय क्षमता कम हो सकती है।**
- **नैतिकता को संहिताबद्ध करने की चुनौतियाँ:** नैतिकता को रोबोट या AI-संचालित सरकारी निर्णयों के लिये स्पष्ट नियमों में रूपांतरित करने का प्रयास एक चुनौतीपूर्ण कार्य के रूप में उजागर किया गया है। मानवीय नैतिकताएँ अत्यंत जटिल प्रकृति रखती हैं और **इन जटिल विचारों को कंप्यूटर निर्देशों में सुसंगत करना कठिन है।**
- **जवाबदेही और पारदर्शिता का अभाव:** AI प्रणाली में कुछ गड़बड़ी होने पर ज़िम्मेदारी का निर्धारण करना कठिन सिद्ध हो सकता है, **विशेष रूप से जब इसमें जटिल एल्गोरिदम और निर्णय लेने की प्रक्रिया शामिल** हो।
    - कई AI प्रणालियों की आंतरिक कार्यप्रणाली प्रायः अपारदर्शी होती है, जिससे यह समझना कठिन हो जाता है कि निर्णय किस प्रकार लिये जा रहे हैं। पारदर्शिता की इस कमी से उपयोगकर्त्ताओं के बीच अविश्वास और संदेह उत्पन्न हो सकता है।
- **सूचना-संपन्न सहमति**: AI सिस्टम का उपयोग संलग्न व्यक्तियों की जानकारी या सहमति के बिना व्यक्तिगत डेटा एकत्र करने और उसका विश्लेषण करने के लिये किया जा सकता है। इससे **सूचना-संपन्न सहमति (informed consent) और निजता के अधिकार (right to privacy) के बारे में चिंताएँ पैदा होती हैं।**

क्या मशीनें या AI नैतिक निर्णय-निर्माता/कृत्रिम नैतिक एजेंट (AMAs) की स्थिति प्राप्त कर सकती हैं?

- कुछ शोध दावा करते हैं कि **मशीनें, एक तरह से, नैतिक दृष्टि से अपने कार्यों के लिये ज़िम्मेदार ठहराई जा सकती हैं** । डार्टमाउथ कॉलेज के प्राध्यापक जेम्स मूर (James Moore) ने **नैतिकता से संबंधित मशीन एजेंटों को चार समूहों में वर्गीकृत किया है**:
    - **नैतिक प्रभाव एजेंट (Ethical Impact Agents):** ये मशीनें रोबोट जॉकी (robot jockeys) की तरह स्वयं नैतिक विकल्प का चयन नहीं करती हैं, लेकिन **उनके कार्यों का नैतिक प्रभाव उत्पन्न होता है।** उदाहरण के लिये, वे किसी खेल (sport) के तरीके को बदल सकती हैं।
    - **अंतर्निहित नैतिक एजेंट (Implicit Ethical Agents):** इन मशीनों में विमानों के ऑटोपायलट की तरह अंतर्निहित सुरक्षा या नैतिक नियम शामिल होते हैं। वे, सक्रिय रूप से यह तय किये बिना कि नैतिक क्या है, **निर्धारित नियमों का पालन करते हैं।**
    - **स्पष्ट नैतिक एजेंट (Explicit Ethical Agents):** ये **निश्चित नियमों से परे जाते हैं।** ये विकल्पों के नैतिक मूल्य का पता लगाने के लिये विशिष्ट तरीकों का उपयोग करते हैं। उदाहरण के लिये, ऐसी प्रणालियाँ जो धन निवेश को सामाजिक उत्तरदायित्व के साथ संतुलित करती हैं।
    - **पूर्ण नैतिक एजेंट (Full Ethical Agents):** ये **मशीनें नैतिक निर्णय ले सकती हैं** और उन्हें समझा सकती हैं। अच्छी नैतिक समझ वाले वयस्क और उन्नत AI इस श्रेणी में आते हैं।

ज़िम्मेदार AI के नैतिक विचार



//

निष्कर्ष

वर्तमान में कई मशीनी पूर्वानुमान निर्णयन प्रक्रिया में मदद करते हैं, लेकिन अंतिम निर्णय अभी भी मानव द्वारा ही लिया जाता है। भविष्य में सरकारें मशीनों को सरल निर्णय लेने की अनुमति दे सकती हैं। लेकिन तब क्या किया जाएगा यदि मशीन द्वारा लिया गया निर्णय गलत या अनैतिक हो? कौन उत्तरदायी होगा? AI प्रणाली उत्तरदायी होगी या वह व्यक्ति/संस्था जिसने AI का निर्माण किया या वह व्यक्ति जिसने इसके डेटा का उपयोग किया?

ये ऐसे कुछ कठिन प्रश्न हैं जिनका सामना विश्व को करना पड़ेगा। मशीनों में नैतिकता का अधिरोपण कठिन है और हर किसी को आगे बढ़ने में सतर्कता बरतनी होगी।

अभ्यास प्रश्न: ''किसी कंप्यूटर को नैतिक बनाने के लिये इसकी प्रोग्रामिंग करना, विश्व-चैंपियन शतरंज खेलने के लिये कंप्यूटर की प्रोग्रामिंग करने से कहीं अधिक कठिन है।'' चर्चा कीजिये।

UPSC सिविल सेवा परीक्षा, विगत वर्ष के प्रश्न

??????????

प्रश्न. विकास की वर्तमान स्थिति में कृत्रिम बुद्धिमत्ता (Artificial Intelligence), निम्नलिखित में से किस कार्य को प्रभावी रूप से कर सकती है? (2020)

1. औद्योगिक इकाइयों में विद्युत की खपत कम करना
2. सार्थक लघु कहानियों और गीतों की रचना
3. रोगों का निदान
4. टेक्स्ट-से-स्पीच (Text-to-Speech) में परिवर्तन
5. विद्युत ऊर्जा का बेतार संचरण

नीचे दिये गए कूट का प्रयोग कर सही उत्तर चुनिये:

(a) केवल 1, 2, 3 और 5
(b) केवल 1, 3 और 4
(c) केवल 2, 4 और 5
(d) 1, 2, 3, 4 और 5

उत्तर: (b)

??????

प्रश्न: "चौथी औद्योगिक क्रांति (डिजिटल क्रांति) के उद्भव ने सरकार के एक अभिन्न अंग के रूप में ई-गवर्नेंस की शुरुआत की है"। चर्चा कीजिये। **(2020)**

PDF Refernece URL: https://www.drishtiias.com/hindi/printpdf/exploring-the-ethical-implications-of-ai